# Decision Memo for Positron Emission Tomography (FDG) for Infection and Inflammation (CAG-00382N)

# **Decision Summary**

Based upon our review CMS has determined that the evidence is inadequate to conclude that FDG PET for chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin improves health outcomes in the Medicare populations and therefore has determined that FDG PET for chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin is not reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act .

We are continuing our national noncoverage of FDG PET for these indications. CMS has also determined that the request for coverage is not appropriate for the Coverage with Evidence Development (CED) paradigm.

Back to Top

## **Decision Memo**

TO: Administrative File: CAG #000382N

PET for Infection and Inflammation

FROM: Steve E. Phurrough, MD, MPA

Director, Coverage and Analysis Group

Louis Jacques, MD Division Director

Stuart Caplan, RN, MAS Lead Analyst

Shamiram R. Feinglass, MD, MPH Lead Medical Officer SUBJECT: Final Coverage Decision Memorandum for Positron Emission

Tomography (PET) for chronic osteomyelitis, infection of hip

arthroplasty and fever of unknown origin.

DATE: March 19, 2008

#### I. Decision

Based upon our review CMS has determined that the evidence is inadequate to conclude that FDG PET for chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin improves health outcomes in the Medicare populations and therefore has determined that FDG PET for chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin is not reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act.

We are continuing our national noncoverage of FDG PET for these indications. CMS has also determined that the request for coverage is not appropriate for the Coverage with Evidence Development (CED) paradigm.

## II. Background

Throughout this memorandum we use the term FDG to refer to 2-[F-18] Fluoro-D-Glucose, also known as fluorodeoxyglucose. We use the term PET to refer to positron emission tomography or to a positron emission tomogram, depending on context. FDG PET refers to PET imaging utilizing FDG as the radioactive tracer.

FDG PET is a minimally-invasive diagnostic imaging procedure used to evaluate glucose metabolism in normal tissue as well as in diseased tissues in conditions such as cancer, ischemic heart disease, and some neurological disorders. FDG is an injected radioactive tracer substance (radionuclide) that gives off sub-atomic particles, known as positrons, as it decays. PET uses a positron camera (tomograph) to measure the decay of radioisotopes such as FDG. The rate of FDG decay provides biochemical information on glucose metabolism in the tissue being studied. As malignancies can cause abnormalities of metabolism and blood flow, FDG PET evaluation may indicate the probable presence or absence of a malignancy based upon observed differences in biologic activity compared to adjacent tissues.

Diagnostic imaging technologies such as x-ray films, computed tomography (CT), and magnetic resonance imaging (MRI) supply information about the anatomic structure of suspected malignancies, primarily their size and location. Clinical imaging of alteration of glucose metabolism within cells is unique to PET technology.

An FDG PET scan can be interpreted based on qualitative and/or semi-quantitative evaluation. Qualitative FDG PET involves making assessments by visually interpreting the scan results. Metabolically active areas of the body "light up" on an FDG PET scan more so than less active areas. Metabolically active areas may include areas of cancer, inflammation and benign cellular activity. Semi-quantitative evaluation uses the glucose metabolic rate of a tumor and, through computer software, determines a numeric value representing the metabolic activity for that tumor. Tumor-to-background ratio is a semi-quantitative method that compares a tumor's glucose uptake to the glucose uptake of surrounding or background tissue. This ratio is reported as the standardized uptake value (SUV) and takes into account such factors as patient weight and injected FDG dosage, as well as the time lapsed from injection to metabolic imaging. There is ongoing debate about the real usefulness of the SUV, especially when comparing results obtained using different PET scanners in different institutions.

Scintigraphy, also known as gamma scintigraphy, is a type of nuclear medicine diagnostic test that obtains a two-dimensional picture of parts of the body. The test is performed by intravenously injecting one of several types of radioisotopes, also known as radioneuclides. The radioisotopes bind with tissues in the body. The radiation emitted by the radioisotope is measured by a specialized machine known as a gamma camera that is placed over the body part being studied. Analyzing the results of the emitted radiation allows the physician to evaluate diseased and healthy bone. Diseased tissue usually emits higher levels of radiation.

Several types of studies and radioisotopes are used in performing scintigraphy. These include white blood cells (WBCs) labeled with indium-111 (In-111) or technetium-99m (Tc-99m) colloidal sulfur, as well as immunoscintigraphy using monoclonal antibodies labeled with technetium-99m. A triple-phase bone scan (TPBS) involves taking a series of gamma camera images at three different times, with several hours between scans.

Current diagnostic tests for osteomyelitis, infections of a revised hip, and fever of unknown origin (FUO)

In chronic osteomyelitis, an infection of the bone, the main disadvantage of any scintigraphic imaging modality using labeled WBCs is that they are unreliable in diagnosing osteomyelitis in the central skeleton, likely because of a loss of sensitivity secondary to low grade chronic infections (Guhlmann 1998).

Serious complications of hip arthroplasty include loosening of the prosthesis and infection which often present as a painful hip. The underlying cause of pain often remains unclear until an intraoperative specimen is examined (Pill 2006). Approximately 8% of all arthroplasties are revisions, of which 70% are for loosening (Reinartz 2005). The current diagnostic tools for proper diagnosis of these conditions still leave clinicians with a diagnostic conundrum when the tests are inconclusive. The most frequent tools are physical examination, clinical laboratory testing (complete blood count [CBC], erythrocyte sedimentation rate [ESR], and C-reactive protein [CRP]), serial CT or MRI, and triple-phase bone scan (TPBS) used with white blood cell scintigraphy (technetium-99m sulfur colloid indium-111 -labeled white blood cell scintigraphy (TcSC-Ind BM/WBC)).

The confirmatory tests are guided hip joint aspiration or surgical sampling, all of which produce a tissue sample that subsequently may be analyzed by histopathology or microbiology. For purposes of this document, "gold standard" refers to these confirmatory tests. TPBS has good sensitivity (ranges 73-100%), but since it has less specificity (30-80%) may lead to more false-positive diagnoses than clinicians would prefer (Mumme 2005). TcSC-Ind BM/WBC is considered to be the imaging modality of choice for the evaluation of infected hips, with a sensitivity range of 60-100% and a specificity range of 58-100% (Pill 2006). Management of infected hip replacements differs substantially from management of hip replacements with aseptic loosening. Revision arthroplasty in aseptic loosening is usually a one-step procedure. Infection in the prosthesis may result in the need for multiple surgical revisions and require antimicrobial therapy (Stumpe 2004).

Conventional radiographs have been reported to be of limited value in the diagnosis of infection as several radiographic findings may be present in both infection and aseptic loosening (Stumpe 2004). However, sequential radiographs may contribute to differentiation because changes occur more quickly in the presence of infection (Stumpe 2004). CT and MRI are also of limited value. Radionuclide studies represent the current imaging method of choice in patients with metallic implants. Because of the high negative predictive value of its results, conventional bone scintigraphy is useful as an initial screening test, but the use of white blood cells labeled with indium-111 (111In), in combination with technetium 99m (99mTc) sulfur colloid marrow imaging, currently provides the highest sensitivity and specificity and has become the method of choice in the assessment of infection in total joint replacements (Stumpe 2004) and likely helpful in FUO and in diagnosing infection in other areas of the body.

## **III. History of Medicare Coverage**

CMS previously reviewed scientific literature and established coverage for many uses of FDG PET. A summary of currently covered PET indications is in the following table. For each indication, there are specific coverage limitations listed in the CMS NCD Manual, Section 220.6.

Currently covered PET indications (FDG unless otherwise noted)

Effective Date	Clinical Condition/Indication	Coverage
March 14, 1995	Myocardial perfusion	Rubidium-82 in coronary artery disease
January 1, 1998	Solitary pulmonary nodule	Characterization
January 1, 1998	Non small cell lung cancer	Initial staging
July 1, 1999	Colorectal cancer	Suggested recurrence with rising CEA
July 1, 1999	Lymphoma	Staging and restaging as alternative to gallium scan
July 1, 1999	Melanoma	Recurrence prior to surgery as alternative to gallium scan
July 1, 2001	Non small cell lung cancer	Diagnosis, staging and restaging
July 1, 2001	Esophageal cancer	Diagnosis, staging and restaging
July 1, 2001	Colorectal cancer	Diagnosis, staging and restaging
July 1, 2001	Lymphoma	Diagnosis, staging, and restaging

Effective Date Clinical Condition/Indication		Coverage
July 1, 2001	Melanoma	Diagnosis, staging and restaging. Non-covered for evaluating regional nodes.
July 1, 2001	Head and neck (excluding central nervous system and thyroid)	Diagnosis, staging and restaging
July 1, 2001	Refractory seizures	Pre-surgical evaluation
July 1, 2001 to September 1, 2002	Myocardial viability	Only following inconclusive SPECT
October 1, 2002	Myocardial viability	Primary or initial diagnosis
October 1, 2002	Breast cancer	Staging, restaging, response to treatment
October 1, 2003	Myocardial perfusion	Ammonia N-13 in coronary artery disease
October 1, 2003	Thyroid cancer	Restaging of recurrent or residual disease
September 15, 2004	Alzheimer's disease and dementia	In CMS-approved clinical trial

Effective Date	Clinical Condition/Indication	Coverage
January 28, 2005	Brain, cervical, ovarian, pancreatic, small cell lung and testicular cancers	Coverage with evidence development
January 28, 2005	All other cancers and indications not previously specified	Coverage with evidence development

## **Current Request**

The requestor		مطالمانيمطاما			£
The reallestor	cinmite that	SUUTIN DE	nationaliv	COVERED	TOT:
THE ICAUCOLOI	Jubility tilat	 JIIOUIU DC	Hauohanv	COVCICA	101.

- 1. Suspected chronic osteomyelitis in patients with:
- (a) previously documented osteomyelitis with suspected recurrence, or
- (b) symptoms of osteomyelitis for more than six weeks (including diabetic foot ulcers)
- 2. Investigation of patients with suspected infection of hip prosthesis. FDG PET would replace bone, leukocyte and/or gallium scintigraphy in the evaluation of these patients.
- 3. Fever of unknown origin in patients with a febrile illness of >3 weeks duration, a temperature of >38.3 degrees Centigrade on at least two occasions and uncertain diagnosis after a thorough history, physical examination and one week of proper investigation.

## **Benefit Category**

Medicare is a defined benefit program. An item or service must fall within a benefit category as a prerequisite to Medicare coverage. § 1812 (Scope of Part A); § 1832 (Scope of Part B) § 1861(s) (Definition of Medical and Other Health Services). FDG PET for the purpose of diagnosis of infection and inflammation is considered to be within the following benefit category: other diagnostic tests §1861(s)(3).

#### IV. Timeline of Recent Activities

June 25, 2007

CMS accepts a formal request for reconsideration of the NCD Manual Section 220.6 to include FDG PET for chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin A tracking sheet was posted on the web site and the initial 30 day public comment period commenced.

July 25, 2007

The initial 30 day public comment period ended. Fifteen comments were received.

December 20, 2007 CMS posts a Proposed Decision Memorandum for a 30-day public comment period.

January 19, 2008 Public comment period ends

#### V. FDA Status

The FDA approved the following uses for FDG F-18 in a Federal Register notice dated March 10, 2000:

"The [FDA] Commissioner has concluded that FDG F-18 injection, when produced under the conditions specified in an approved application, can be found to be safe and effective in FDG-PET imaging in patients with [coronary artery disease] CAD and left ventricular dysfunction, when used together with myocardial perfusion imaging, for the identification of left ventricular myocardium with residual glucose metabolism and reversible loss of systolic function, as discussed in section III.A.1 and III.A.2 of this document. The Commissioner also has concluded that FDG F-18 injection, when produced under the conditions specified in an approved application, can be found to be safe and effective in FDG-PET imaging for assessment of abnormal glucose metabolism to assist in the evaluation of malignancy in patients with known or suspected abnormalities found by other testing modalities or in patients with an existing diagnosis of cancer, as discussed in section III.A.1 and III.A.3 of this document. In addition, manufacturers of FDG F-18 injection and sodium fluoride F-18 injection may rely on prior agency determinations of the safety and effectiveness of these drugs for certain epilepsy-related and bone imaging indications, respectively, in submitting either 505(b)(2) applications or [amended new drug applications] ANDAs for these drugs and indications."

The FDA approval language cited above indicates that FDG F-18 is not currently approved by the FDA to assist in the diagnosis of infection. Therefore this use of FDG PET imaging would represent an off-label use.

## VI. General Methodological Principles

When making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member. The critical appraisal of the evidence enables us to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for Medicare beneficiaries. An improved health outcome is one of several considerations in determining whether an item or service is reasonable and necessary.

A detailed account of the methodological principles of study design that are used to assess the relevant literature on a therapeutic or diagnostic item or service for specific conditions can be found in Appendix A. In general, features of clinical studies that improve quality and decrease bias include the selection of a clinically relevant cohort, the consistent use of a single good reference standard, and the blinding of readers of the index test, and reference test results.

Public comment sometimes cites the published clinical evidence and gives CMS useful information. Public comments that give information on unpublished evidence such as the results of individual practitioners or patients are less rigorous and therefore less useful for making a coverage determination. CMS uses the initial public comments to inform its proposed decision. CMS responds in detail to the public comments on a proposed decision when issuing the final decision memorandum.

#### VII. Evidence

A. Introduction
We are providing a summary of the evidence that we considered during our review.
A diagnostic test must provide information that is used by the treating physician to appropriately guide the management of the patient's specific medical problem (42 CFR 410.32).
As a diagnostic test, FDG PET would not be expected to directly change health outcomes. Rather, a diagnostic test affects health outcomes through changes in disease management brought about by physician actions taken in response to test results. Such actions may include decisions to treat or withhold treatment, to choose one treatment modality over another, or to choose a different dose or duration of the same treatment. To some extent the usefulness of a test result is constrained by the available treatment options. In addressing the questions below, one of the factors we consider is whether there is sufficient evidence that the incremental information derived from FDG PET leads to improved treatment of the disease in question by causing physicians to prescribe a different treatment than they would have prescribed without access to the test results.
Outcomes of interest for a diagnostic test are not limited to determining its accuracy but also include beneficial or adverse clinical effects, such as changes in management due to test findings or preferably, improved health outcomes for Medicare beneficiaries. Ideally, we would see evidence that the systematic incorporation of FDG PET results into a treatment algorithm leads treating physicians to prescribe different treatment than they would otherwise have prescribed, and that patients whose treatment is changed by test results achieve better long term outcomes.
B. Discussion of evidence reviewed
Questions
<ol> <li>How does the diagnostic test performance of FDG PET compare to bone, leukocyte and/or gallium scintigraphy with respect to the following clinical situations:</li> </ol>

	a. b. c.	Chronic osteomyelitis in patients with previously documented osteomyelitis with suspected recurrence or symptoms of osteomyelitis for more than 6 weeks? Infection associated with hip arthroplasty? Fever of unknown origin where febrile illness is: 1) greater than three weeks duration; 2) a temperature of greater than 38.3 degrees centigrade occurs on at least two occasions, and 3) diagnosis is uncertain after a thorough history, physical examination and one week of appropriate investigations?
2.		evidence sufficient to conclude that FDG PET can replace bone, leukocyte and/or gallium scintigraphy for dications listed in Question 1?
3.	manag	evidence sufficient to conclude that FDG PET for the indications listed in Question 1 changes patient gement or improves patient oriented outcomes when compared to bone, leukocyte and/or gallium graphy?
2. Exte	ernal te	echnology assessments
An ext	ernal T <i>i</i>	A was not commissioned.
3. Inte	rnal ted	chnology assessments
		ned an extensive literature search utilizing PubMed for new randomized controlled trials (RCTs). The rch was limited to the English language and specific to the human population using search terms:
FDG P FDG P	PET, ost PET, sci PET, hip	ver of unknown origin steomyelitis sintigraphy, comparative study o arthroplasty dium-111

Printed on 7/22/2011. Page 11 of 35

FDG PET, technetium-99m FDG PET, immunoscintigraphy FDG PET, triple phase bone scan FDG PET, gamma scintigraphy

FDG PET, monoclonal antigranulocyte antibody FDG PET, infection
The current request for coverage of FDG PET for the various infection and inflammation indications includes 13 documents, citations of which are provided in the references section. A search in the Cochrane Library failed to reveal any systemic reviews evaluating the use of FDG PET for the requested indications.
This review will be restricted to studies with specified outcomes. Any studies with less than 20 participants were not considered for this NCA. In addition, clinical review articles that were submitted were not considered (Vos 2006, Zhuang H (2004)).
Subsequent to the issuance of the proposed decision, CMS conducted an additional literature search using the above criteria. No additional relevant articles were found.
Since there are three distinct diagnostic conditions being addressed in this NCA, the evidence will be reviewed by indication. We believe that restating each of these individually in the Analysis section will make for cumbersome reading. Thus, for the convenience of the reader each piece of evidence is followed by a brief summary of its limitations as noted either by the author(s) or by CMS.
Individual Study Results:
CHRONIC OSTEOMYELITIS
Guhlman, et al. 1998: This blinded, prospective case series compared FDG PET vs. immunoscintigraphy (IS) with 99m Tc-labeled monoclonal antigranulocyte antibodies in 51 patients with suspected chronic osteomyelitis. The final diagnosis was determined by histopathology or culture (n=31) or by biopsy and clinical follow-up over 2 years (n=20). Thirty-six peripheral and 15 central skeletal infections (12 women, 39 men; age range 22-81; mean age 48.5) were diagnosed. All patients also had a 99mTc-MDP bone scan.
Of 51 patients, 28 had osteomyelitis and 23 did not.

	Sensitivity (%)	Specificity (%)	Accuracy (%)
TPBS (all)	86-92	77-82	82-88
PET (all)	97-100	95	96
TPBS (p)	89-93	82-88	86-92
PET (p)	95-100	95	95-97
TPBS (c)	80-90	60	73-80
PET (c)	90-100	100	93-100

all=all patients; p=peripheral skeleton; c=central skeleton

Limitations: A sample size of 15 is very small. Study authors concluded that a larger study is needed to confirm these preliminary results.

<u>Kalicke</u>, et al. 2000: This was a case series of 21 patients suspected of having acute or chronic osteomyelitis or inflammatory spondylitis (12 men, 9 women; age 33-78). Of these, 15 underwent surgery and FDG PET was correlated with histopathology. The other six were excluded from further evaluation because of no histopathology. Only 11 underwent bone scintigraphy (the comparator).

Limitations: The sample size is too small to show significance.

Printed on 7/22/2011. Page 13 of 35

<u>De Winter, et al. 2001:</u> This blinded, prospective case series evaluated chronic musculoskeletal infections, including chronic osteomyelitis, spondylodiscitis or post- joint arthroplasty, in 60 patients (33 in central skeleton, 27 in peripheral skeleton). The age range was 13-75. Final diagnosis was based on histopathology or culture in only 18 of 60 patients. In the other 42 patients, the determination of whether an infection was present was based on clinical findings after at least six months follow-up.

	Sensitivity (%)	Specificity (%)	Accuracy (%)
Overall	100	88	93
Central Skeleton	100	90	94
Peripheral Skeleton	100	86	93

Authors conclude: "Larger series are needed to define the role of FDG PET in the evaluation of acute osteomyelitis...accurate and less expensive techniques for the detection of acute osteomyelitis are available and the added value of FDG PET is likely to be limited."

Limitations: This study is limited by the use of a heterogeneous selection of infection sites (spine, hip, femur); low level of histopathologic confirmation; and the limits inherent to case series methodology. Additionally, FDG PET was not compared to bone, leukocyte and/or gallium scintigraphy, which are the best available tests and used to guide patient management.

Zhuang, et al. 2000. In this retrospective case series of 22 patients suspected of chronic osteomyelitis at a variety of sites (tibia, spine, femur, pelvis, maxilla, feet), FDG PET was evaluated compared to histopathology. FDG PET sensitivity was 100% and specificity was 87.5%. Results were as follows: 6 true positive, 16 true negative, 2 false positive and no false negative.

Limitations: There were no data tables presented and this was a retrospective case series; as such, no clear conclusion may be drawn. FDG PET was not compared to bone, leukocyte and/or gallium scintigraphy, which are the best available tests and used to guide patient management. The study authors conclude that positive results can be caused by inflammation in the bone or surrounding soft tissues as a result of other causes (surgery, etc). Hence, a positive scan may be inconclusive regarding osteomyelitis. A negative scan may be helpful.

<u>Schiesser M, 2003</u>: This prospective case series of 29 FDG PET scans in patients suspected of having osteomyelitis as a result of a metallic implant (not joint replacement) was blinded and compared FDG PET scans to surgical specimens in 26 cases. Ages ranged from 18-86.

Localization	Sensitivity (%)	Specificity (%)
Peripheral Skeleton (n=20)	100	87.5
Central Skeleton (n=9)	100	90
All cases (n=29)	100	93.3

Limitations of this study were that not all cases were compared to surgical samples, it is a case series and the sample size of the subgroups is too small to show significance regarding the test performance of FDG PET scans. Finally, the cases were limited to post-trauma surgical interventions with metallic hardware; hence generalizability to the larger Medicare population is questionable.

Meller J, et al. 2002: This blinded, prospective case series compared FDG PET to indium-111-labelled WBCs to diagnose chronic bacterial osteomyelitis. Of 30 cases, only 18 were compared to histology or culture for TPBS and 19 for FDA PET. Age range was 24-72. Thirty cases had TPBS. Performance statistics are limited to those with histological confirmation; hence the sample size is 18 for TPBS and 19 for FDG PET.

	True Positive	True Negative	False Positive	False Negative
TPBS (n=18)	2	8	1	7
FDG PET (n=19)	9	10	0	0

Limitations of this study are the case-series design and that histological confirmation was not obtained for all samples.

<u>Termaat, et al. 2006:</u> This meta-analysis of several imaging modalities for the assessment of chronic osteomyelitis only evaluated four FDG PET studies and FDG PET was not the sole focus of the meta-analysis. The heterogeneity of the articles used for the analysis limit the applicability of this publication to the questions at hand.

## HIP

All studies submitted by the requestor were considered, though two will not be discussed in detail as they all involved sample sizes too small to provide meaningful results: Manthey 2002: (n=14) and Vanquickenborne 2003 (n=17).

Mumme, et al. 2005: This case series of FDG PET vs. TPBS in loose hip arthroplasty examined 50 patients (70 hips; 50 with pain, 20 asymptomatic) 31 women and 19 men. The average age was 68.7 (range 42-86). The final diagnosis was made during operative revision in all 50 symptomatic hips. The remaining asymptomatic patients (20 hips) were followed clinically. FDG PET performed better than TPBS.

	True positive	True negative	False positive	False negative	Sensitivity (%)	Specificity (%)
TPBSª	37	21	6	6	86	76
TPBSb	31	21	9	9	78	70
FDG PETª	43	22	2	3	93	92
FDG PET <sup>b</sup>	42	22	2	4	91	92

<sup>&</sup>lt;sup>a</sup>Detecting pathological processes

Limitations: This was a case series and these types of studies are a weaker methodological design than RCTs and as such are limited in their ability to distinguish useful from useless or even harmful therapy. It is unclear from the published manuscript if the study was blinded, and only 50 cases were compared to surgical sampling (a gold standard).

Pill, et al. 2006: This blinded case series compares PET with technetium-99m sulfur colloid indium-111-labeled white blood cell scintigraphy (TcSC-Ind BM/WBC). Results were verified by histology or patient outcome when surgery was not indicated. Eighty-nine patients scheduled to undergo revision hip arthroplasty (92 hips; 6 bilateral) were included. An additional 36 patients who had undergone hip arthroplasty and were without postoperative pain were recruited as controls.

bDifferentiation between septic and aseptic loosening

Test	N	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
FDG PET	92	95.2	93	80	93
TcSc-Ind BM/WBC	51	50	95.1	41.7	88.6
ESR	92	14.3	47.9	10	65.3
CRP	92	52.4	26.8	20.7	65.5
CRP+ESR	92	93	89	75	99
WBC	92	23.8	97.2	71.4	81.2
Joint Aspiration	16	0	93.8	0	100

PPV - positive predictive value NPV - negative predictive value

Limitations: This was a case series which limits the generalizability of the results and may bias the results. Only 51 patients had TcSc-Ind BM/WBC. Since fewer patients had TcSc-Ind BM/WBC or were compared to histology, selection bias is introduced into the evidence.

Printed on 7/22/2011. Page 18 of 35

Reinartz, et al. 2005: This blinded case series of 63 patients (92 hips, 31 female) compares FDG PET vs. TPBS for loosening and/or infection of hip arthroplasty. The reported sensitivity, specificity and accuracy of FDG PET were greater than for bone scan. The mean age was 68. For revised hips, final diagnosis was based on intra-operative findings as well as histopathology and microbiology.

	True positive	True negative	False positive	False negative	Sensitivity (%)	Specificity (%)
TPBSª	27	51	7	7	79	88
TPBSb	17	51	16	8	68	76
FDG PETª	32	56	2	2	94	97
FDG PET <sup>b</sup>	31	56	3	2	94	95

<sup>&</sup>lt;sup>a</sup>Detecting pathological processes

<sup>b</sup>Differentiation between septic and aseptic loosening

TPBS: triple-phase bone scanning

Limitations: Not all cases were compared to surgical findings (gold standard). Also, this study is a case series.

<u>Stumpe</u>, et al. 2004: This blinded, case series of 35 people (23 women, mean age of women was 64; men was 71) examined FDG PET vs. conventional radiology and TPBS for infected hip arthroplasty. Results showed that FDG PET is more specific but less sensitive than conventional imaging. The final diagnosis was verified by microbiology and intraoperative findings.

Diagnostic Performance of FDG PET, Conventional Radiography, and Three-Phase Bone Scintigraphy for infected hip

	Sensitivity (%)	Specificity (%)	Accuracy (%)
FDG PET	22-33	81-85	69
Conventional Radiography	78-89	50-65	60-69
ГРВЅ	44-56	88-92	80

The authors conclude that their data "suggest that FDG PET as an infection imaging modality offers no benefit in addition to three-phase bone scintigraphy in patients with prosthetic joint replacement... PET performed similarly to three -phase bone scintigraphy. PET was more specific but less sensitive than conventional radiography for the diagnosis of infection."

Limitations: The authors note that the prevalence of infection was low, which may affect sensitivity. In addition, this was a nonrandomized case series.

Delank, et al. 2006: This blinded, prospective case series of 27 patients with a painful hip or knee (n = 22 hips) compared FDG PET to triple-phase bone scan (TPBS). Results were verified by intraoperative histopathology and microbiology. Loosening was correctly identified in 76.4% of FDG PET scans and 75% in TPBS. Sensitivity is 100% for identifying septic inflammation, but only 45.5% in inflammation due to increased abrasion and aseptic foreign-body reactions. Reliable differentiation between septic (bacterial-induced) and aseptic (abrasion induced) inflammation is not possible.

The authors note that "mechanical loosening cannot be sufficiently identified using [FDG PET]."

Limitations to this study are that this is a case series and the sample size was small.

Zhuang, et al. 2001: This prospective case series examined painful hips and knees post arthroplasty. FDG PET was performed in thirty-eight hips in this combined hip and knee study. The comparator was surgical sampling or clinical follow-up. Sensitivity for infection was 90%, specificity was 89.3%, and accuracy 89.5%. This was a preliminary study.

Limitations: This study is small (authors note that this was a preliminary study), and it was unclear how many FDG PET results were compared to surgical sampling (a gold standard). Finally, FDG PET was not compared to triple-phase bone scan, which is one of the best available tests and is used in most of the hip literature.

## **FEVER OF UNKNOWN ORIGIN (FUO)**

<u>Bleeker-Rover</u>, et al. 2006: This prospective study of 70 patients with FUO presented a "structured diagnostic protocol" of two tiers, one of which included FDG PET. Not all patients received both tiers of testing. Final diagnosis was by biopsy, positive serology or positive culture. The mean age of participants was 53 (26-87). Thirty-three scans were deemed clinically helpful but were more useful in patients with continuous fever. Diagnosis was never based on FDG PET alone. The cause of FUO was not determined in 50% of cases. Thirty-three scans were abnormal.

	True positive	True negative	False positive	False negative	Sensitivity (%)	Specificity (%)
FDG PET	23	34	10	2	88	77

The 34 true negatives included 26 patients without a final diagnosis who had a normal PET scan and were classified as true negative. Results were confirmed by biopsy, microbiology, serology, or patient follow-up. The authors compared PET to abdominal CT in 60 patients.

Management change was interpreted based on the diagnostic protocol and the number of unnecessary tests avoided by correct FDG PET results. Hence, authors conclude that 70% of PET scans were clinically helpful and ultimately contributed to the ultimate diagnosis in 33% of patients.

Limitations: Twenty-six patients with a normal PET scan without a final diagnosis were classified as having true negative results, but this is a significant reporting bias. Since the results were inconclusive (the true source of the fever was never found and the patient remained febrile), then these scans should have been reported as indeterminate and analyzed as such. The authors note: "Calculation of sensitivity and specificity of FDG PET in patients with FUO is difficult, because a final diagnosis is not established in all patients. When additional diagnostic procedures, performed according to the diagnostic protocol, are negative and long-term follow-up does not reveal a diagnosis, it is probably legitimate to presume that focal infectious disease, inflammation or malignancy is not the cause of the symptoms in these patients."

CMS does not agree with this rationale as it is speculative and it is not supported by the literature. The goal of the test is to find the cause of fever. Just because one was not found (a negative FDG PET scan) does not mean one does not exist and that other testing modalities would not find the cause. To classify the negative scans in this case as true negative results in misclassification. Finally, this rationale leads to significant bias towards FDG PET scans; this classification system will overestimate the usefulness of PET.

FDG PET was not useful in examining the lower legs because of the way investigators performed the scans. The authors also note that in patients with negative FDG PET, a variety of diseases were found that could not be diagnosed with FDG PET.

Finally, this was a case series and FDG PET was not compared to bone, leukocyte and/or gallium scintigraphy which are the best available tests.

<u>Blockmans</u>, et al. 2001: This case series of 58 consecutive patients with FUO compared FDG PET to gallium scintigraphy (n=40). Final diagnosis was established in 38 patients. The goal was to evaluate PET as a "second step technique" in a diagnostic protocol. No protocol was used for FUO workup; instead guidelines were available for first-step examinations (laboratory testing, history and physical examination). No patient ages were provided. The authors report that 24 of 46 abnormal scans correctly identified the fever source and were clinically helpful (though it is unclear how the correct diagnoses were made). When compared to gallium scintigraphy, FDG PET was deemed useful in 35% of patients and gallium in only 25%.

The authors conclude that an abnormal FDG PET scan could "be used as an indication for further intelligent testing (CT, MRI, biopsy of involved foci), but not for invasive procedures such as exploratory laparotomy"

Limitations: All patients had FDG PET, but only 40 had gallium scintigraphy because of the limited availability of gallium at the study site. This has decreased the sample size of the study by 18 patients or 31%. In addition, the generalizability is limited because patient demographics were not reported and it is unclear how the final diagnoses were made.

<u>Jaruskova</u>, et al. 2006: This was a retrospective case series of 124 patients with FUO (94), or fever due to a suspected infection of a joint or vascular prosthesis (30). The goal of the study was to "determine the frequency which FDG PET or PET/CT could help establish a diagnosis in a group of symptomatic patients in whom a variety of previous examinations had failed to yield definitive results." The goal was not to compare PET to any other modality. Clinical confirmation was not obtained on all patients. Neither age nor sex of the patients was reported. Fifty-one patients had a confirmatory test for their diagnosis and FDG PET was contributory in reaching the final diagnosis in 42 of these patients.

Limitations: The retrospective case series nature of the study, lack of follow-up information on patients, and lack of patient demographics all limit generalizability. Additionally, it is unclear what confirmatory tests were done or what the management changes were secondary to FDG PET scan results. FDG PET was not compared to bone, leukocyte and/or gallium scintigraphy, which are the best available tests and used to guide patient management.

## 4. MedCAC

A Medicare Evidence Development and Coverage Advisory Committee (MedCAC) meeting was not convened on this issue.

## 5. Evidence-based guidelines

We are unaware of any evidence-based guidelines for the use of FDG PET for the diagnosis of chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin.

## 6. Professional Society Position Statements

The Society of Nuclear Medicine and the American College of Radiology jointly submitted a letter that supported coverage of FDG PET for the requested indications. Five references were cited in this letter, none of which were unique from those submitted by the requestor. The Academy of Molecular Imaging also submitted a letter expressing support of the NCD request. Thirteen citations were provided, each of which was also submitted by the requestor.

7. Expert Opinion
We have not solicited or received any expert opinions on the use of FDG PET for the requested indications.
8. Public Comments
Initial Comment Period: June 25, 2007 - July 25, 2007
CMS received a total of 15 comments during the first public comment period. All of the comments supported coverage of FDG PET for the requested indications. Eight of the comments were from physicians. The remaining seven comments were from nuclear medicine technicians, billing coders and others. GE Healthcare voiced its support for the request but cautioned that CMS should not require that FDG PET replace conventional nuclear medicine studies. Articles submitted with these public comments were not unique to those submitted by the requestor or identified by CMS during its literature review.
Comment period on the proposed decision: December 20, 2007 – January 19, 2008
CMS received six comments on the proposed decision. Two of the comments were from physicians, one of which was from the requestor. We note that of the commenters who self-identified as physicians, none stated clinical practice as orthopedic surgeons or infectious disease specialists.
Comment GE Healthcare, the Academy of Molecular Imaging and the requestor expressed support for the request and suggested that CMS cover the requested indications under the Coverage with Evidence Development (CED) paradigm.
Response We have thoughtfully considered the use of CED for this indication but have determined that the current evidence is insufficient to support CED at this time.

#### Comment

It was also suggested that CMS encourage the National Institutes of Health (NIH) to undertake a clinical trial of PET for infection.

## Response

While we have no authority to determine the NIH's research priorities, we agree that further clinical trials are needed on this topic.

## Comment

One requestor posited that PET for the requested indications would alleviate the risk of handling blood.

## Response

It is not clear from the evidence that PET would alleviate the risks associated with handling blood. FDG is administered intravenously. Even if a PET scan were suggestive of an infection it would still be necessary to obtain blood cultures to guide antimicrobial therapy.

## Comment

Two other commenters expressed opinions in favor of the request based on clinical experience. No articles were submitted with these public comments.

#### Response

While we appreciate the commenters' personal experience, we must accord it less evidentiary weight than evidence from more methodologically rigorous clinical trials.

#### VIII. CMS Analysis

National coverage determinations (NCDs) are determinations by the Secretary with respect to whether or not a particular item or service is covered nationally under title XVIII of the Social Security Act § 1869(f)(1)(B). In order to be covered by Medicare, an item or service must fall within one or more benefit categories contained within Part A or Part B, and must not be otherwise excluded from coverage. Moreover, with limited exceptions the expenses incurred for items or services must be "reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member." §1862(a) (1) (A). This section presents the agency's evaluation of the evidence considered and conclusions reached for the assessment question:

#### General

Changes in patient management are brought about by physician actions taken in response to test results. Such actions may include decisions to treat or withhold treatment, to choose one treatment modality over another, or to choose a different dose or duration of the same treatment. 42 CFR 410.32(a) states in part, "...diagnostic tests must be ordered by the physician who is treating the beneficiary, that is, the physician who furnishes a consultation or treats a beneficiary for a specific medical problem and who uses the results in the management of the beneficiary's specific medical problem."

We generally assign greater weight to evidence produced by randomized clinical trial (RCT) design since this methodology provides the strongest evidence of causal linkages (see Appendix A). However, there were no RCTs upon which to base this coverage determination. Only case series were available and these types of studies are a weaker methodological design than RCTs. Nonrandomized studies of efficacy are limited in their ability to distinguish useful from useless or even harmful therapy. As is noted in the following sections and detailed in the Evidence section, all the reviewed studies suffer from the impact of confounding that could occur between variables studied, as well as other threats to internal validity (e.g., selection bias, reliability of measures and procedures).

Reported study results are most consistent about the potential for FDG PET to be useful for patient management decisions for any of the three indications when pain has persisted for a minimum of 12 months post-surgical intervention. This criterion is important because the post-surgical inflammation may lead to false-positive results from a PET scan regardless of the underlying diagnosis. Twelve months seems to be the consistent time frame used by authors to reduce this likelihood of a false positive result (Mumme 2005).

As addressed in the Evidence section, the few studies that even had a comparator were not consistent in the standards they used for comparison with FDG PET. In addition, comparisons could not be made across studies regarding how FDG PET test statistics compared to the gold standards of guided aspiration or surgical sampling and the resulting histological or microbiological results. In addition, comparisons could not be made across studies regarding how FDG PET test statistics compared to bone, leukocyte and/or gallium scintigraphy, which are currently the best available tests and used to guide patient management. Some studies were more suggestive than others, but the balance of the studies could not be compared.

We will address each question individually below.

1. How does the diagnostic test performance of FDG PET compare to bone, leukocyte and/or gallium scintigraphy with respect to the following clinical situations:

a.	Chronic osteomyelitis in patients with previously documented osteomyelitis with suspected recurrence or symptoms of osteomyelitis for more than 6 weeks?
b.	Infection associated with hip arthroplasty?
C.	Fever of unknown origin where febrile illness is: 1) greater than three weeks duration; 2) a temperature of greater than 38.3 degrees centigrade occurs on at least two occasions, and 3) diagnosis is uncertain after a thorough history, physical examination and one week of appropriate investigations?
<b>5</b> 1	
availat vhen ( CMS c	was a general lack of consistency in comparators used to establish FDG PET as a viable alternative to the best ble tests of bone, leukocyte and/or gallium scintigraphy. Furthermore, there was a general lack of reported data CMS examined the individual studies. Given this lack of consistency and lack or reported data in the literature, oncludes that good diagnostic test performance of FDG PET for any of the three indications is not demonstrated. iled analysis follows.
OSTE	OMYELITIS
and for arger is ample overifuit on the compaction of the compacti	cally for chronic osteomyelitis, only Guhlman, et al (1998) compared FDG PET to scintigraphy and bone scan and only moderate improvement in sensitivity (from 86-92% for TPBS to 97-100% for FDG PET), though a slightly improvement in specificity (77-82% for TPBS to 95% for FDG PET). However, given the limitations of this study e size of 15) no conclusion of benefit may be made. Furthermore, study authors note that a larger study is needed by the conclusions. The rest of the studies tried to establish the test statistics for FDG PET compared to histology, iology or patient follow-up. These studies were fraught with limitations including small sample size; lack of rison to bone, leukocyte and/or gallium scintigraphy, which are currently the best available tests and used to guide management; and they did not consistently compare all cases to histology, microbiology or other tissue sampling ty, which would be the diagnostic gold standards. Thus, we are unable to conclude that there is a benefit for are beneficiaries. This leads us to conclude that FDG PET is not equal to bone, leukocyte and/or gallium raphy.

If an infection in a post-surgical hip could be ruled out by a negative FDG PET scan that has a high sensitivity, the operative consequences of revision of an infected hip might be avoided. Additionally, a prominent abrasion-mediated inflammatory reaction, which would eventually lead to loosening, could be identified using FDG PET well before such a reaction would be recognizable radiologically or scintigraphically (Delank 2006). The test could make it possible to decide whether close radiological observation is necessary or whether, in certain cases with advanced osteolytic changes, early revision should be contemplated. Unfortunately, the current literature does not support this hypothesis.

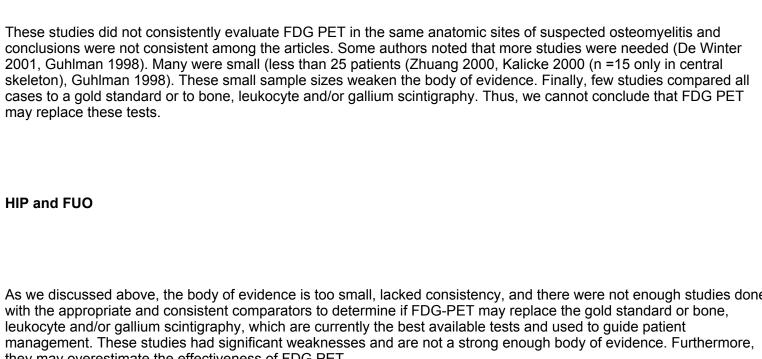
Though the literature relating to hip was of better quality than for the other indications, it was still limited by small sample sizes and the nature of the study design which, in all cases, was a case series. Most of these studies did compare FDG PET to at least one of the best available tests, usually TPBS, but the results were mixed. Stumpe, et al. (2004) found that TPBS performed better than FDG PET in all cases and suggested that there was no benefit to the use of FDG PET. On the other hand, Reinartz, et al. (2005) found FDG PET test performance to be significantly better that TPBS. There is clear disagreement in the literature. Given the heterogeneity of the study results and their comparators, the test performance of FDG PET does not compare favorably with testing modalities currently used as the standard of care. Furthermore, FDG PET was not always compared to a tissue diagnosis; some cases were followed "clinically." This lack of consistency in test performance results leads CMS to conclude that FDG PET is not equal to the gold standard or to the best available tests.

## **FEVER OF UNKNOWN ORIGIN (FUO)**

Only three studies addressed FUO and only one of them (Blockmans, et al. (2001)) was structured to compare the test performance of FDG PET to one of the best available diagnostic tests (gallium scintigraphy). Furthermore, the Blockmans study was designed only to determine FDG PET scan use as a second step in a diagnostic protocol. This study was limited because only 40 of 58 patients were compared to gallium scintigraphy and no test performance statistics were presented. There was only a report that 24 of 46 abnormal PET scans correctly identified the fever source. However, patient demographics were not presented and it is unclear how the final diagnosis was made. Hence, CMS concludes that FDG PET is not equal to the gold standard or to bone, leukocyte and/or gallium scintigraphy in relation to correctly identifying the source of an FUO.

2. Is the evidence sufficient to conclude that FDG PET can replace bone, leukocyte and/or gallium scintigraphy for the indications listed in Question 1?

## **OSTEOMYELITIS**



As we discussed above, the body of evidence is too small, lacked consistency, and there were not enough studies done they may overestimate the effectiveness of FDG PET.

We cannot determine from the evidence that FDG PET compares favorably with the gold standards or the best available tests because the body of literature suffers from inconsistency and small sample size. Thus, we cannot conclude that FDG PET may replace these tests.

3. Is the evidence sufficient to conclude that FDG PET for the indications listed in Question 1 changes patient management or improves patient oriented outcomes when compared to bone, leukocyte and/or gallium scintigraphy?

#### CHRONIC OSTEOMYELITIS

Many of these studies did not have bone, leukocyte and/or gallium scintigraphy as a comparator, nor were all scans compared to a gold standard (surgical tissue sample). Furthermore, the literature included a variety of anatomical sites to study. The inconsistent use of a comparator or a gold standard and the mix of anatomical sites of study are significant limitations. The literature seems consistent in noting that bone scans are less helpful in the central skeleton and that the role of FDG-PET might be for this scenario. However, the results are inconsistent and the sample sizes too small to be conclusive. Thus, for chronic osteomyelitis CMS cannot conclude that FDG PET changes patient management or improves patient oriented outcomes.

HIP

Given that the clinical management of an infected hip replacement vs. loosening of a replacement differ significantly, it would be important to have a diagnostic tool that would better differentiate between septic and aseptic loosening, be less invasive than biopsy, and with lower risk than tagged WBC scans.

Authors suggest that PET scans may have a role in management of painful hip replacements as PET is not invasive and may have similar test result statistics to the standard practice armamentarium. However, the study results regarding the usefulness of FDG PET for differentiating periprosthetic infection and aseptic loosening have been mixed as is detailed in the individual results section and in question 1. Thus, for the differentiation of infected hip replacement versus loosening of the replacement CMS cannot find that FDGPET changes patient management or improves patient oriented outcomes.

## **FUO**

Limitations of FDG PET scans remain—they are unreliable in distinguishing infection vs. neoplastic processes. Though this might be a benefit for diagnosing the site of an FUO, since both conditions cause FUO, it still leaves a large diagnostic conundrum.

Of the evidence reviewed to answer the question of how FDG PET may change management in the patient with FUO, the studies were conducted only to see where FDG PET might fit into the diagnostic work-up. While this is an attempt to determine how management might change using this technology, it does not accomplish this task in an appropriate, evidence rich way. The literature base was small and limited by a lack of consistent comparison to a gold standard or the best available diagnostic tests. Given these weaknesses, we find that FDG PET does not provide useful information that can be used for patient management or to improve patient oriented outcomes.

#### Summary

Our review of the evidence showed significant limitations in the body of evidence inlcuding heterogenaeity of cases, lack of consistent comparators, small sample sizes, missing data (or lack of data reporting), and poor methodologic structure, with one study in particular (Bleeker-Rover 2006) leading to an overestimate of the test performace of FDG PET. In addition, most studies were case series which are useful for hypothesis generation, but do not show causality.

Although two of the clinical conditions that are the subject of this decision (chronic osteomyelitis and hip replacement) are commonly managed by orthopedic surgeons we note the marked paucity of expressed interest on this issue by practicing orthopedic surgeons or their professional societies. Similarly we note the lack of expressed interest from those physicians, generally infectious disease specialists, who would routinely be asked to consult in cases of fever of unknown origin. This leads us to reasonably determine that the interest in the use of PET for these indications is narrow and does not apparently include the physicians who routinely manage the care of beneficiaries who have these conditions.

The evidence reviewed does not support a finding that FDG PET reliably informs physician management or improves patient centered outcomes for these indications. Thus we have determined that FDG PET is not reasonable and necessary for these indications.

#### **CED**

We did not propose CED in the proposed decision memorandum. Rather, we requested that the public provide information as to the specific types of studies that would be appropriate under CED. Though we received comment regarding coverage with the collection of additional clinical data, we did not received a proposal for a clinical study that would fulfill the requirements of Coverage with Study Participation (CSP). Thus, we conclude that CED is not appropriate at this time.

## IX. Conclusion

Based upon our review CMS has determined that the evidence is inadequate to conclude that FDG PET for chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin improves health outcomes in the Medicare populations and therefore has determined that FDG PET for chronic osteomyelitis, infection of hip arthroplasty and fever of unknown origin is not reasonable and necessary under Section 1862(a)(1)(A) of the Social Security Act.

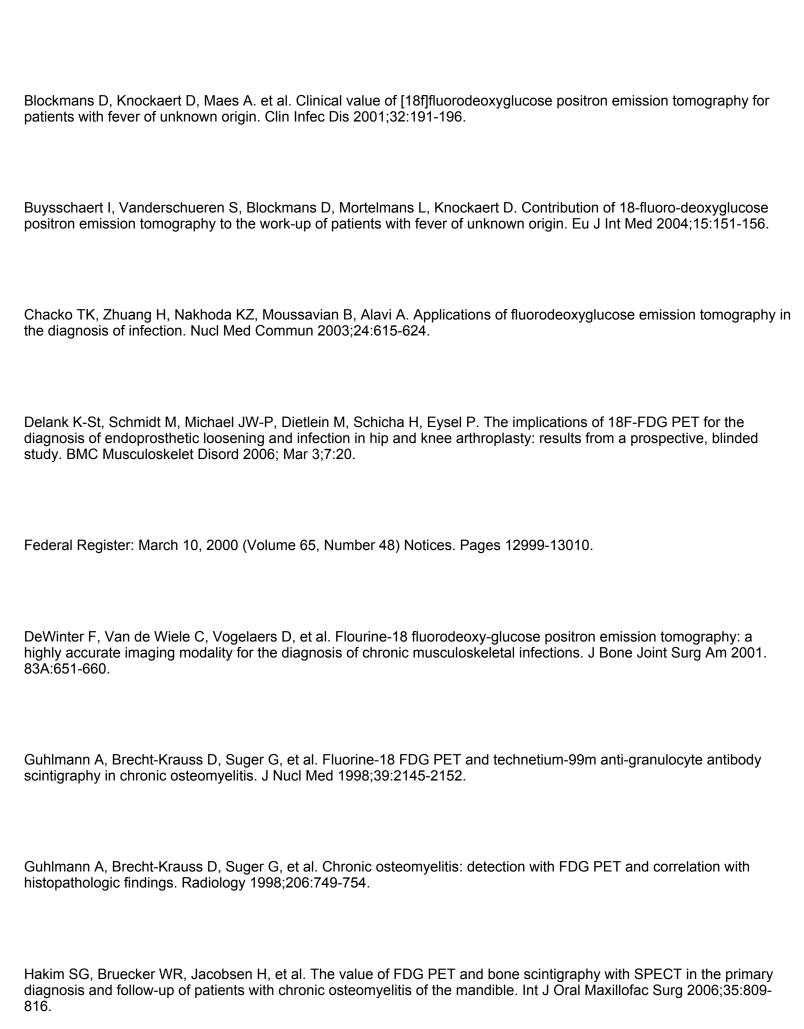
We are continuing our national noncoverage of FDG PET for these indications. CMS has also determined that the request for coverage is not appropriate for the Coverage with Evidence Development (CED) paradigm.

Back to Top

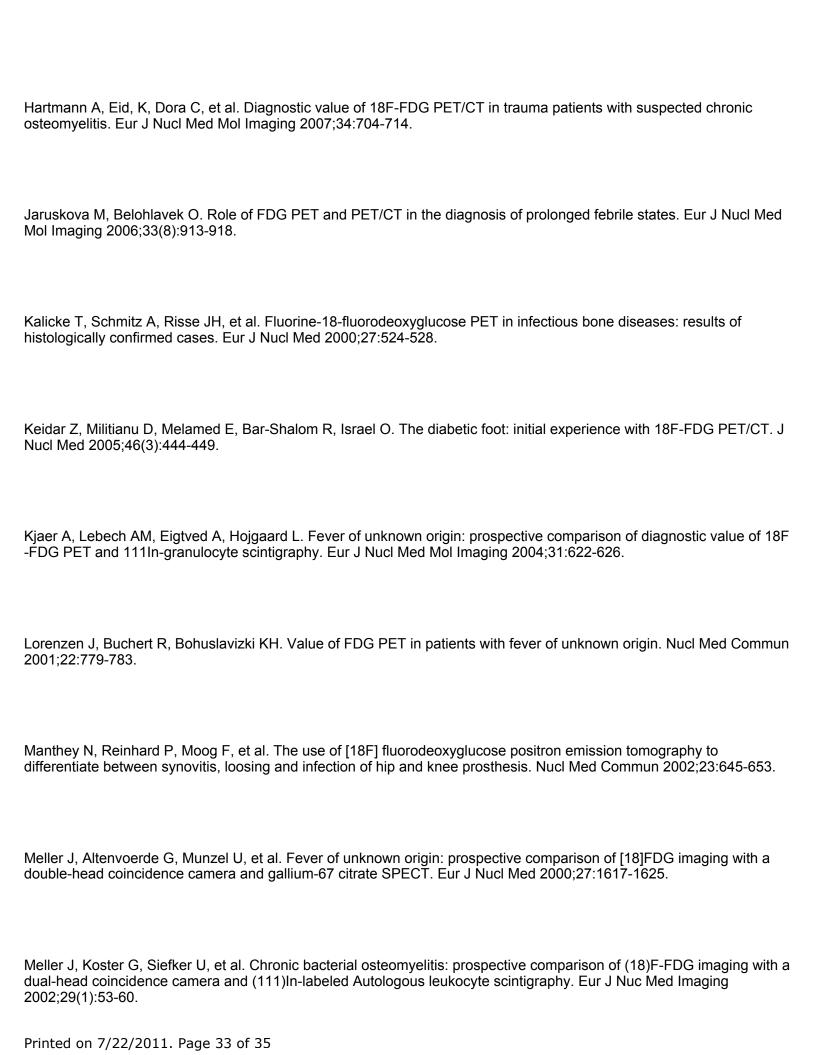
## **Bibliography**

Bleeker-Rovers CP, Vos FJ, Mudde AH, et al. A prospective multi-centre study of the value of FDG PET as part of a structured diagnostic protocol in patients with fever of unknown origin. Eur J Nucl Med Mol Imaging 2006; Epub.

Printed on 7/22/2011. Page 31 of 35



Printed on 7/22/2011. Page 32 of 35



Meller J, Sahlmann CO, Scheel AK. 18F-FDG PET and PET/CT in fever of unknown origin. J Mucl Med 2007;48:35-45.
Mumme T, Reinartz P, Alfer J, Muller-Rath R, Buell U, Wirtz DC. Diagnostic values of positron emission tomography versus triple-phase bone scan in hip arthroplasty loosening. Arch Orthop Trauma Surg 2005;125:322-329.
Pill SG, Parvizi J, Tang, PH, et al. Comparison of fluorodeoxyglucose positron emission tomography and 111Indium-white blood cell imaging in the diagnosis of periprosthetic infection of the hip. The Journal of Arthroplasty 2006;21(6):suppl. 2.
Reinartz P, Mumme T, Hermanns B. Radionuclide imaging of the painful hip arthroplasty: positron-emission tomography versus triple-phase bone scanning. J Bone Joint Surg Br 2005;87:465-470.
Stumpe KD, Notzli HP, Zanetti M, et al. FDG PET for the differentiation of infection and aseptic loosening in total hip replacements: comparison with conventional radiography and three-phase bone scintigraphy. Radiology 2004;231:333-341.
Termaat MF, Raijmakers PGHM, Scholten HG, et al. The accuracy of diagnostic imaging for the assessment of chronic osteomyelitis: a systematic review and meta-analysis. J Bone Joint Surg Am 2005;87:2464:2471.
Vanquickenborne B, Maes A, Nuyts J, et al. The value of 18FDG PET for the detection of infected hip prosthesis. Eur J Nucl Mol Imaging 2003;30:705-715.
Vos FJ, Bleeker-Rovers CP, Corstens FHM, Kullberg BJ, Oyen WJG. FDG PET for imaging of non-osseous infection and inflammation. Q J Nucl Med Mol Imaging 2006;50:121-130.
Zhuang H, Duarte PS, Pourdehnad M, et al. The promising role of 18F-FDG PET in detecting infected lower limb prosthesis implants. J Nucl Med 2001;42:44-48.

Zhuang H, Duarte PS, Pourdehnad M, Shnier D, Alavi A. Exclusion of osteomyelitis with F-18-fluorodeoxyglucose positron emission tomographic imaging. Clin Nucl Med 2000;25(4):281-284.

Zhuang H, Yu JQ, Alavi, A. Applications of fluorodeoxyglucose-PET imaging in detection of infection and inflammation and other benign disorders. Radiol Clin North Am 2005;43:121-134.

Back to Top